# Transformación no lineal

En esta clase vamos a seguir trabajando los métodos de escalamiento.

Resumen:
Los métodos lineales se usan para distribuciones normales o que pueden ser Gaussianas o de forma uniforme, ahí la idea es que el escalamiento comprime los datos a un rango que es adecuado para los algoritmos de Machine Learning.

## ¿Qué debo hacer cuando los datos no cumplen con la distribución simétrica?

Cuando la distribución no es uniforme y hay sesgamiento en los datos. Haremos algo que en matemáticas es muy común y es tenemos un problema que no es común, lo que hago es transformarlo a un problema que se como resolver, son 2 pasos.

- Transformar el problema que no se resolver a uno que tengo conocimiento de ello.
- Le aplico el método de resolución que ya me se.

Ahora aplicado a nuestro caso es:

- Tomamos los datos que no están distribuidos de forma simétrica, les aplicamos una transformación. Esta transformación los deja distribuidos de forma simétrica.
- Después de eso, aplicamos los escalamientos lineales.

# Transformaciones no lineales

### ¿Por que usarlos?

Porque hay datos fuertemente sesgados y no simétricos.

### Tipos de transformaciones no lineales

Es decir usamos aquellas funciones no lineales de grado 2 que tenemos en matemáticas, para buscar un comportamiento más homogéneo de los datos y hacerlos simétricos.

- Logaritmos
- Sigmoideos
- Polinomiales
- etc

### ¿Cuando usarlos?

Antes del escalamiento lineal.

## Recuerda

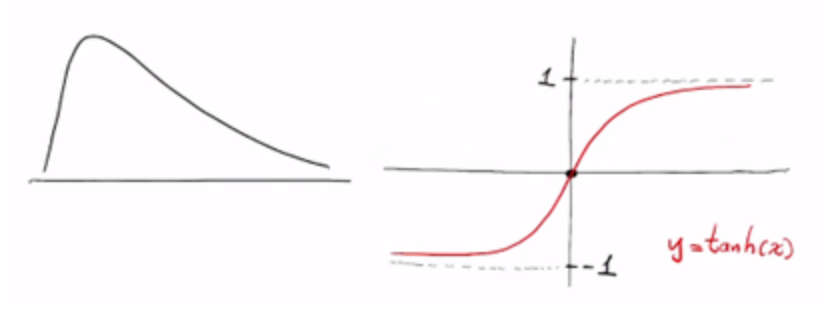
Al final siempre tenemos que aplicar el escalamiento lineal, siempre y cuando garantices que la distribución es simétrica.

En este caso hay un bloque adicional para cuando surgen este tipo de cuestiones. Así que diremos:

Tengo mi **pipeline** dividido en 2 bloques; cuando no tengo una simetría en mis datos, hago este proceso para hacerlo simétrico y así aplicar después el escalamiento lineal para que los datos queden en el rango adecuado de Machine Learning.

## Manos a la obra

$$X_s = tanh(x)$$



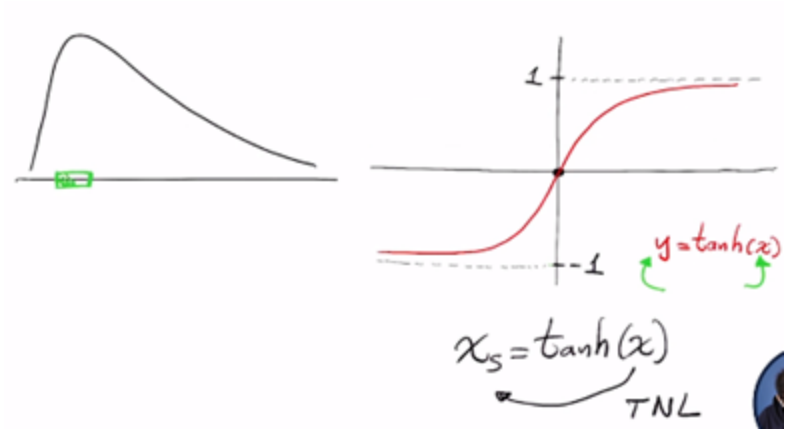

Vamos a ver que tenemos una distribución como por ejemplo a la de la gráfica anterior, veremos que le hace la función no lineal. La idea o el truco de la tangente hiperbólica, es que su intervalo esta entre $[-1, 1]$ y ahí lo interesante es ver cuando $x$ es interpretado como los **datos originales** y los **datos transformados** como $y$

$$y = tanh(x)$$

Entonces nuestra transformación quedaría como:

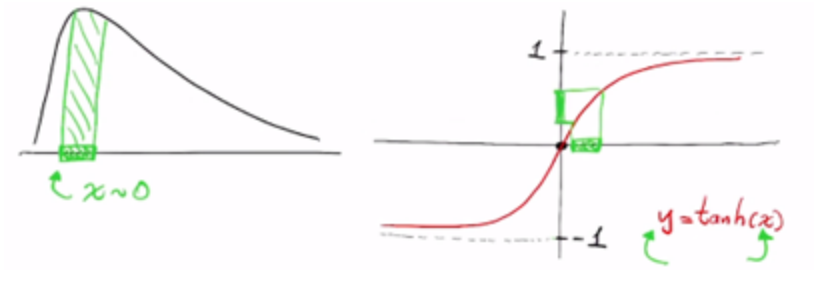
$$X_s = tanh(x)$$

Donde $X_s = Transformación\ No\ Lineal\ (TNL)$.



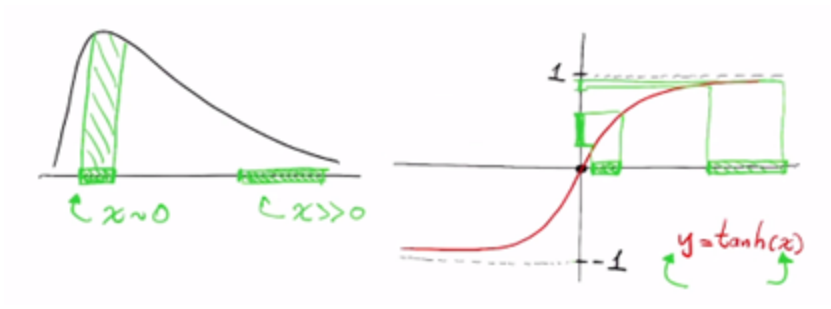

### ¿Que pasa con datos $x \approx 0$?

El resultado de la variable transformada es un intervalo que en general puede resultar un poco mas ancho que el intervalo original, como se puede observar en la gráfica de la tangente hiperbólica.




### ¿Que pasa con datos cercanos a valores outliers?

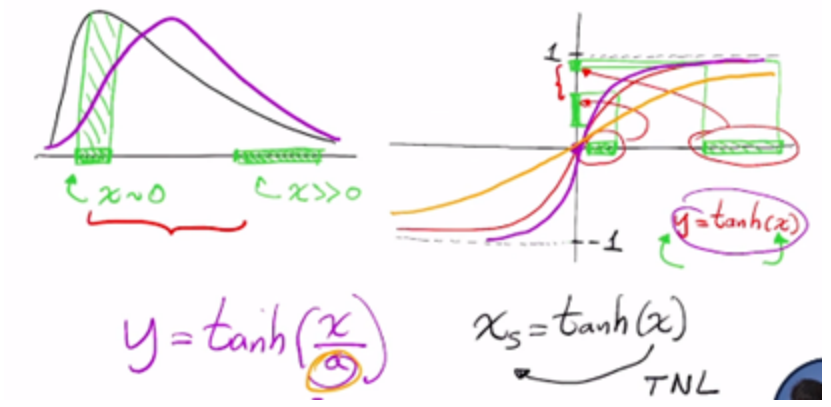
Estos datos son mucho mas grandes $x >>> 0$, entonces cuando yo los mapeo por medio de una **tangente hiperbólica**, lo que hace es mapear los datos a un rango aproximado a 1.




Entonces el efecto de esto es impresionante, porque lo que esta pasando aquí es que antes tenía un intervalo de datos muy distantes entre ellos y que ahora en el mapeo con **tangente hiperbólica** están muy cercanos entre ellos, eso es un **escalamiento** también solo que no sucede de manera homogénea. El resultado gráfico se verá mas claro en las siguientes clases en un notebook de Python; es que la distribución que obtengo como resultado de esos datos podría tener un sesgo reducido y podría ser, sino bien uniforme podría estar un poco mas simétrica, que es justamente lo que deseamos.

### ¿Cómo puedo ajustar el parámetro de simetría?

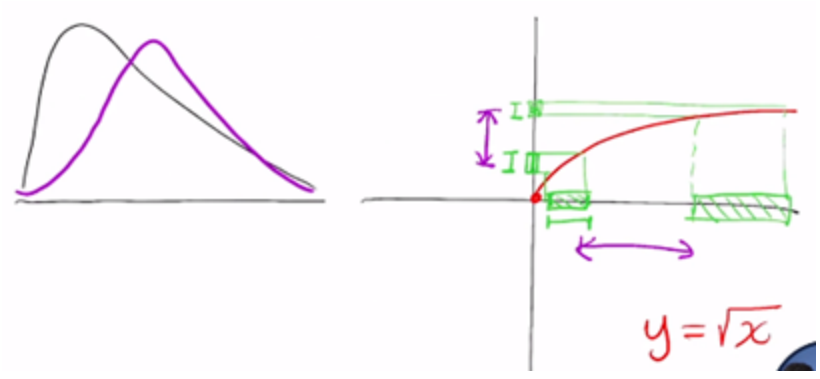
Eso depende de la función que usa para transformar, en este caso de $tanh(x)$. Las **tangentes hiperbólicas** tienen la característica de que yo puedo dividirla por un parámetro $a$ y dependiendo de los valores del parámetro $a$, la pendiente de la **tangente hiperbólica** puede ser más pronunciada (gráfica morada) o menos pronunciada o suave(gráfica naranja).
Gracias a los diferentes valores del parámetro $a$ yo podría ajustar la **deformación de mis datos** para que queden mas simétricos.




$$X_s = tanh\left[\frac{X}{a}\right]$$

$$X_s = \sqrt{x}$$

Al igual que en el caso anterior, en los valores cercanos a $x = 0$ la distribución de la variable $x$ es mas estrecha y para valores $X >>> 0$ los valores tienen a ser más uniformes (en este caso podríamos hablar de valores outliers).




En algunos casos esto también se puede expresar de manera polynomial por que $y = \sqrt{x} = x^{\frac{1}{2}}$, entonces yo podría usar otro tipo de función polinómica.

$$X_s = X_n \text{ donde } n > 0$$

Es decir $n$ podría ser cualquier numero **positivo** y dependiendo de esto vamos a tener diferentes efectos en la distribución.

> El tema es que el abanico de posibilidades de todas las trasformaciones no lineales es enorme.

## Resumen:

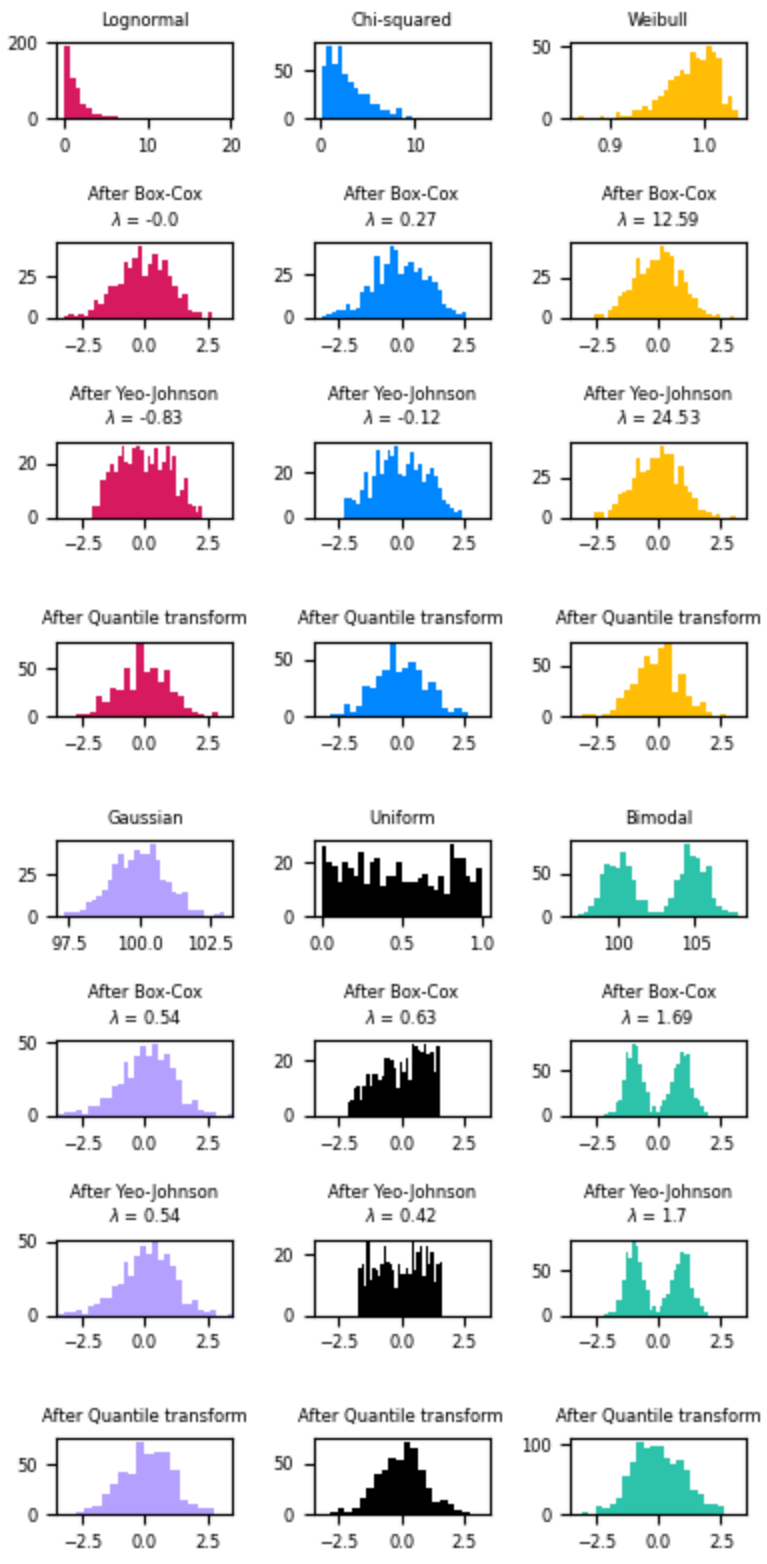
El ejercicio nos va a decir , cual es la función que mas se ajusta.

Para el caso de el profesor prefiere:

- **Tangente hiperbólica**
- **Sigmoide** Porque son muy parecidas y tiene esa propiedad de que:

> Los rangos cercanos a $0$ los expanden mientras que los outliers los contraen y los acercan al origen, haciendo que la distribución sea mas homogénea.

## Extra:




- Normalization
- 3Blue1Brown